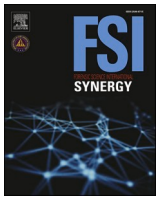
Forensic Science International: Synergy 13 (2026) 100714

Contents lists available at ScienceDirect

Forensic Science International: Synergy

journal homepage: www.sciencedirect.com/journal/forensic-science-international-synergy



Sally Clark: Trial and error by Judge and Jury

Charles Feeny [a], Carolina Stamboulid [b,*]

[a] Complete Counsel, Liverpool, United Kingdom
[b] Episteme Scientific Consulting, Liverpool, United Kingdom

ARTICLE INFO

Keywords:
Sally Clark
Expert evidence
Statistical evidence
Sudden infant death syndrome
Forensic reasoning
Evidential reliability

ABSTRACT

Sally Clark's 1999 conviction for the murder of her two infant children represented one of the most conspicuous miscarriages of justice in modern English criminal law. A central component of the prosecution case was statistical evidence presented by Professor Sir Roy Meadow, which suggested that the probability of two sudden infant deaths occurring naturally within the same family was approximately one in 73 million. This paper examines three methodological limitations that affect the evidential value of that estimate. First, the family characteristics used to reduce the estimated probability of natural death were not considered in relation to the probability of the alternative explanation, thereby distorting comparative inference. Second, sudden infant death syndrome is a diagnosis of exclusion; population frequency data derived from such a category cannot determine causation in an individual case and are contingent on the limits of postmortem investigation. Third, the probability estimate was presented without reference to the likelihood of competing explanations, limiting its interpretive value. The paper further argues that the significance of the Clark case lies not only in the arithmetical error in the one in 73 million figure but in the broader evidential process by which probabilistic evidence was admitted, framed and interpreted within a criminal trial. The paper also examines the empirical and unstructured approach of English criminal courts to expert evidence. It considers the extent to which subsequent developments, including the Criminal Practice Directions 2023, respond to these difficulties, and compares that position with the contrasting procedural framework associated with the Daubert standard in the United States. The analysis situates the Clark case within broader issues concerning the interpretation, contextualisation and scrutiny of expert evidence in legal proceedings.

1. Introduction

On 9 November 1999, Sally Clark, a 35-year-old solicitor, was convicted of the murder of her two children [1]. Her first son, Christopher, was found lifeless on the evening of 13 December 1996. He was 11 weeks old. Her second son, Harry, was similarly found on 26 January 1998. He was nine weeks old. In both cases, Sally Clark had been alone with each child immediately prior to their death. Both bodies were subject to postmortem examination carried out by a Home Office pathologist, Dr Williams. After the death of Christopher, he concluded that the child had died as a result of a respiratory tract infection. He revised this conclusion after carrying out a postmortem examination on Harry, when he believed that he had found retinal haemorrhages consistent with a shaking injury. He therefore reported that Harry had died as a result of shaking. In the light of this conclusion, he revisited his assessment of Christopher, reporting that Christopher had died as a result of smothering. In each case, he described findings which he indicated were consistent with physical abuse [1,2].

Sally Clark was convicted by a jury majority of ten to two [1]. She appealed her convictions. Her first appeal in October 2000 was rejected by the Court of Appeal. Her second appeal in January 2003 was allowed. It was allowed because it had come to light that, in his postmortem report on Harry, Dr Williams had not mentioned the existence of a microbiology report which showed significant evidence of *Staphylococcus aureus* bacteria [2]. The significance of that finding lay in the fact that *S. aureus* is a recognised human pathogen, capable of causing severe invasive disease and death [3]. The undisclosed microbiology report therefore provided a plausible natural or infective line of explanation for Harry's death. This made it at least plausible, indeed likely according to experts called by the defence, that he had died from natural causes rather than deliberate harm. Dr Williams had neither discussed the contents of the microbiological tests nor revealed the report's existence

This article is part of a special issue entitled: The Science behind Wrongful Convictions published in Forensic Science International: Synergy.

* Corresponding author.
E-mail address: carolina.stamboulid@episteme-consulting.co.uk (C. Stamboulid).

https://doi.org/10.1016/j.fsisyn.2026.100714
Received 28 April 2026; Received in revised form 24 June 2026; Accepted 26 June 2026
Available online 30 June 2026

during the course of Sally Clark's prosecution and trial. He was discredited as an expert and an examiner. Since much of the prosecution expert evidence depended upon the integrity of his postmortem findings, the evidential basis for guilt had collapsed. Although Sally Clark was released from prison after the second appeal, she never recovered from the trauma of losing two children and then being wrongly imprisoned. She died in 2007, aged 42, from accidental acute alcohol intoxication [4, 5].

The successful second appeal therefore turned primarily on the undisclosed microbiological evidence and the consequent collapse of confidence in the postmortem findings [2]. The statistical evidence presented by Professor Sir Roy Meadow is not treated in this paper as the sole or formal legal cause of the wrongful conviction but as an illustration of a distinct and continuing problem in criminal adjudication. Its significance lies in the way technically flawed expert evidence can alter the evidential environment in which criminal fact-finders assess guilt.

The jury who convicted Sally Clark would not have known about the undisclosed microbiology report nor of the full extent of the deficiencies in Dr Williams' evidence. They did, however, hear evidence from Professor Meadow, a paediatrician called by the prosecution, who stated that the probability of two sudden infant deaths occurring naturally within a family such as the Clarks was approximately one in 73 million. That figure later became one of the most widely criticised aspects of the case. It attracted concern because of the assumptions on which it rested, the way it was capable of being understood by a jury, and the absence of a proper comparative framework against which its evidential value could be assessed [6–9].

The statistical evidence in the Clark case has already been the subject of substantial criticism. Those critiques have addressed the presentation of statistical evidence by non-statisticians in criminal trials, the assumptions made about a jury's ability to understand and discount such evidence, the need to compare the probability of the evidence under competing hypotheses and the forensic misuse of epidemiological probability where population-derived estimates are applied to an individual case without adequate foundation [7,10,11]. The present concern is not only that the one in 73 million figure was statistically flawed but that it entered the trial as an apparently authoritative expression of rarity without the comparative and methodological structure required for it to have reliable evidential meaning.

The analysis brings those strands together within a broader account of evidential failure, extending beyond the arithmetical error in the one in 73 million figure to distinguish the statistical invalidity of the calculation, the use of sudden infant death syndrome as a residual diagnostic category, the absence of comparative probabilistic reasoning, the risk of the prosecutor's fallacy, the boundary between paediatric and statistical expertise, and the adequacy of legal mechanisms for excluding or controlling unreliable expert evidence.

The article finally considers the contemporary significance of the Clark case. The precise combination of errors that occurred in 1999 is unlikely to recur in the same form, yet the case continues to reveal a more general vulnerability in the reception of expert evidence, including the possibility of analogous failures where weaknesses in technical reasoning are not properly identified before the evidence reaches the jury.

2. Expert evidence before *Clark*: qualification, reliability and the common law approach

In the judgment on the second appeal, the Court dealt only briefly with the issues in relation to Professor Meadow's evidence. They had already determined that the appeal should be allowed because of the problems with Dr Williams' evidence. The Court considered in relation to Professor Meadow's evidence that *"the wisest course would have been to exclude it altogether"* (paragraph 177) [2]. The significance of this wording is that the Court did not identify a structured admissibility principle by which the evidence should have been excluded. The issue was expressed retrospectively as one of prudence rather than as the application of a defined reliability threshold.

2.1. The common law approach to expert admissibility

The most serious criminal cases and some lesser cases where the accused has a right of election are determined in the Crown Court by a judge and jury. The process of trial by judge and jury has a long history stretching back into the 12th century [12–14]. It emerged as a melding of two different strands of determination in criminal cases; that is the jury, which could reasonably be described as a northern European system of village elders, and the development of the role of the king's judges. The system entrusts application of the law to the judge, therefore relying on their wisdom, and determination of the facts to the jury, relying on their good sense. Whilst the criminal law in England has now been the subject of considerable statutory intervention, the guiding principles still remain largely empirical in origin and development. So, by 1999 the courts had not developed any structured approach with clearly identifiable principles relating to the admissibility of expert evidence.

2.2. R v Robb and methodological disagreement as weight rather than admissibility

In *R v Robb* [1991] [15], Lord Justice Bingham stated, in relation to expert evidence:

> "The text which the English common law has developed is characteristically pragmatic".

The reference to the common law as "characteristically pragmatic" amounted to a tacit acknowledgement that there was no structured or principled approach to the assessment of expert reliability. The case illustrates the nature of that pragmatism. The defendant, *Robb*, appealed his conviction in circumstances in which the prosecution had relied upon Dr Baldwin as an expert in phonetics. Whilst Dr Baldwin was suitably qualified, having a PhD in phonetics and a university position as a lecturer in the subject, the method he used for voice identification was not accepted by the majority of experts in his field. He conceded that *"the great weight of informed opinion, including the world leaders in the field"* would not approach the exercise upon which the evidence was directed in the same way. He also accepted that a number of Western European countries, including Germany, specifically did not permit the approach which he had taken [15].

The defendant, who had been convicted on evidence which included Dr Baldwin's identification of his voice in a phone call, appealed on the basis that the evidence should not have been admitted. Lord Justice Bingham indicated that the Court *"have not found this an entirely easy question"*. He conceded that there was a risk in a situation of this nature that evidence could be put before a jury inappropriately, stating that *"a defendant cannot fairly be asked to meet evidence of opinion given by a quack, a charlatan or an enthusiastic amateur"* [15].

Ultimately, however, the Court concluded that the evidence was admissible and that it was for the jury to decide what weight to place upon it, having regard to the concessions made about its shortcomings. The significance of *Robb* is therefore that, once the expert was treated as sufficiently qualified, even serious methodological disagreement within the relevant field could be treated as a matter of weight for the jury rather than admissibility for the judge. That approach is central to the present discussion because the Clark case exposes what may happen when technically complex but flawed expert evidence reaches the jury without first being subjected to adequate methodological scrutiny.

3. Evidence of Professor Sir Roy Meadow

3.1. Meadow's Law and the framing of recurrent infant death

There could be no doubt that Professor Meadow had ample qualification to give evidence as a paediatrician. He was a professor of paediatrics. He was recognised as an expert with particular interest in child abuse. He regularly gave evidence in the Family Courts, which determined allegations relevant to custody and care of children. He, like a number of other experts, had taken a particular interest in sudden infant death syndrome (SIDS), more colloquially known as cot death. The concept of sudden infant death syndrome involved a situation, as occurred on two occasions in Sally Clark's case, of a child being found dead with no immediate indication as to how the child had died [16]. Professor Meadow and a number of other paediatricians were of the view that a significant number of these cases represented deliberate acts by a parent; that is, homicide. He encapsulated this view in what became known as Meadow's Law: *"One sudden death in a family's a tragedy, two is suspicious and three is murder unless proven otherwise"* [16]. This proposition was a heuristic rather than an evidential principle. Its relevance to the Clark case lies in the way it framed recurrent unexplained infant death as intrinsically suspicious, forming part of the evidential background against which Professor Meadow's statistical evidence was later presented [6,10,17].

3.2. The CESDI estimate and the one in 73 million figure

In his initial report served by the prosecution prior to trial, Professor Meadow indicated that the likelihood of the same family experiencing two cases of SIDS was in the order of one in one million. Shortly prior to trial, he amended this assessment to one in 73 million. He did this in circumstances where the prosecution had encountered a very significant difficulty in proving their case. At a meeting of experts prior to trial, the prosecution experts had agreed that there was no evidence of retinal haemorrhage in the findings at Harry's postmortem; therefore, no evidence that he had been shaken. The prosecution therefore changed their case to indicate that Harry had been smothered, as was already their case in relation to Christopher.

Professor Meadow's revised estimate of the likelihood of two SIDS deaths in the Clark family was explained by him through his awareness of an unpublished study. This was the Confidential Enquiry into Stillbirths and Deaths in Infancy (CESDI) study on sudden unexplained death in infancy (SUDI) [18]. Professor Meadow was writing the preface to the study. It was available in draft at the trial but only quoted from selectively. The analysis of the prevalence of SIDS in that study took account of a number of factors, many of which related to socioeconomic status, such as smoking or whether there was a waged income in the household. Unsurprisingly, this study indicated a much lower rate of SIDS in affluent households. It was on this basis, namely the asserted rarity of SIDS in such households, that Professor Meadow increased his estimation from one in one million to one in 73 million.

3.3. Reference class and comparator error

The difficulty with this adjustment was not only that it made natural death appear exceptionally rare in affluent families but that it did so by relying on family characteristics whose relevance was considered only in relation to SIDS. Affluence, a stable relationship and non-smoking may reduce the estimated probability of SIDS but those same features were also likely to reduce the likelihood of homicide. A family which is statistically low-risk for one explanation would also be statistically low-risk for the alternative explanation [6]. The probative value of the calculation was therefore undermined because the characteristics used to make natural death appear rarer were not assessed in relation to the competing hypothesis [6,19].

3.4. The prosecutor's fallacy and jury interpretation

At trial, there was no attempt to exclude this evidence as inadmissible but only a limited challenge to its value and credibility. In the first appeal, again there was no ground of appeal that the evidence was inadmissible. There was a submission that the jury had been misdirected in relation to this evidence. The defence lawyers relied upon an earlier decision of the Court of Appeal in *R v Doheny and Adams* [20], in which the Court had warned against what was described as *"the prosecutor's fallacy"*. This fallacy was identified as the mistaken conversion of a specific conclusion from statistical evidence into an assessment of the likelihood of the defendant's guilt or innocence. It was argued that the jury may have understood the evidence as meaning that there was a one in 73 million chance of Sally Clark being innocent. Remarks by the prosecution and in the judge's summary to the jury were consistent with this impression [2].

The Court of Appeal accepted *"that there is some substance to the criticism that the Judge appeared to endorse the prosecution's erroneous approach in this particular"* [1]. They described the one in 73 million figure as being *"of no help"* and *"merely a distraction"*. They therefore found that, whilst the figure was a distraction, the jury had not been distracted by it. They dismissed the appeal on the basis of their own analysis of all the evidence in the case, concluding that, in the absence of the misleading statistical evidence, the remaining evidence would nonetheless be considered overwhelming. They stated, *"the point on statistics was of minimal significance and there is no possibility of the jury having been misled so as to reach verdicts that they might not otherwise have reached"* [1]. They made this statement even though the Appeal Court was informed that the one in 73 million figure had been widely reported as a prominent feature of the case in the press.

3.5. The independence assumption

By the time of the second appeal, the evidence of Professor Meadow had been subject to trenchant criticism from the Royal Statistical Society in a news release of 23 October 2001 [6]. The contents of the news release were substantially repeated in a letter from the President of the Royal Statistical Society to the Lord Chancellor of 23 January 2002 [9]. In that letter they described the approach of squaring the likelihood of one death to produce the one in 73 million figure as being *"statistically invalid"*. It would only be valid if SIDS cases arose independently within families, an assumption that would need to be justified empirically [8,9]. No such empirical justification was provided in the case and there were also strong a priori reasons for doubting that the assumption could be correct. Genetic, familial and environmental factors are shared within families and persist across siblings; genetic architecture, intrauterine environment, postnatal care patterns and socioeconomic conditions are not independent variables [6,10]. In those circumstances, the probability of a second death could not properly be derived simply by squaring the probability of a first [6,8,9].

The CESDI study, which provided the basis for Professor Meadow's figure, itself contained data which bore on this question, though its significance in this respect was not examined at trial, nor analysed in the published literature until some years afterwards. Hill, writing in 2004, demonstrated that among families who had already lost one child to SIDS, the risk of a second was between six and ten times greater than for families with no such history [6]. Drawing on the CESDI dataset and corroborating that finding against the Care of Next Infant Programme and a Norwegian whole-population study [6]. This analysis was not available at the time of trial. However, the assumption of independence, which the Royal Statistical Society identified as unjustified, could have been interrogated by reference to the same dataset on which Professor Meadow relied [8].

In their letter, the Society stated that the *"Society does not tolerate doctors making serious clinical errors because it is widely understood that such errors could mean the difference between life and death. The case of R v*

Sally Clark is one example of a medical expert witness making a serious statistical error, one which may have had a profound effect on the outcome of the case". They then indicated that statistical evidence should only be given in court by *"appropriately qualified statistical experts"* [9].

The discussion of the statistical evidence in the second appeal was brief and not determinative of outcome [2]. It mirrored a number of the well-known criticisms of the evidence. The Court observed that the limitations of the study were not referred to by Professor Meadow. These included the purpose of the study, which was not to inform diagnosis in individual cases but rather to direct resources. The authors of the study had specifically cautioned against making any assumptions from the occurrence of two SIDS cases in the same family [2,18]. Whilst the Court did indicate that it would have been 'wisest' to exclude the evidence, they did not identify a detailed admissibility principle by which that conclusion should have been reached. They did not discuss whether it was appropriate for Professor Meadow, as a paediatrician, to give evidence that depended upon statistical reasoning.

Expert witnesses in many scientific disciplines have to engage with statistical evidence. There is therefore room for debate as to whether the advice of the Royal Statistical Society, that statistical evidence should only be given by fully qualified statisticians, should be followed rigidly in every case. The more limited point is that where the probative value of expert evidence depends upon statistical assumptions, population estimates or recurrence probabilities, the court requires some mechanism for ensuring that those assumptions have been properly formulated and tested.

In considering the unreliable nature of this evidence, the focus has been upon the major point that such occurrences were treated as if they were random and did not reflect any genetic, familial or environmental factors. Such factors must be at work, as evidenced by the fact that a number of the witnesses at trial were aware of cases of double SIDS in families. Professor Meadow's estimation of there being one such case every hundred years was not consistent with the clinical experience of a number of the expert witnesses.

3.6. Probability without a competing hypothesis

The criticism of the one in 73 million figure has understandably focused on why it was wrong. It is also necessary to consider what it did not establish. A probability attached to one possible explanation has little evidential weight unless it is considered alongside the probability of the competing explanation. Professor Meadow told the jury how rare two natural cot deaths were said to be. He did not address how rare two murders in such a family would be. The figure was therefore presented in isolation, inviting the jury to attach weight to it when this was not justified. That omission was not incidental because the trial process imposed no obligation on an expert witness to place a statistical figure within a comparative framework. Dawid's analysis of Clark makes the same point in evidential terms. If background evidence of double death rates due to SIDS or other natural causes was relevant, then background evidence of double death rates due to murder was also relevant; to present one without some assessment of the other was irrational and prejudicial [11].

3.7. SIDS as a residual diagnostic category

A further difficulty lies in the nature of SIDS itself. SIDS is not a positive diagnosis based upon identified pathological criteria. It is a residual label applied when no sufficient explanation for death can be found [21,22]. Applying population frequency data derived from a residual diagnostic category to the question of what caused a particular death mistakes the nature of what such data can show. The figure relied upon did not represent the probability of two established natural deaths. It represented the frequency with which two deaths had been placed within an unexplained category in a particular dataset, a figure dependent on what investigators were able to identify through postmortem examination [6,18]. In a case where the postmortem process itself was later shown to be flawed, any inference drawn from it was correspondingly unreliable [2].

One of the defence medical experts, Professor Berry, indicated that he would describe the cause of death as unascertained and stated: *"it means that the child's death may have been natural but without explanation, perhaps what the jury knows as* cot *death. It might be that the child died unnaturally but I cannot find out why or it might be that the child died of natural disease or that I am not clever enough to diagnose and recognise and that is why examination of children found suddenly and unexpectedly dead has to be done more thoroughly and more carefully perhaps than any other type of postmortem examination"* [2].

Against this background, it is not coincidental that the miscarriage of justice arose from inaccurate and incomplete analysis of postmortem findings. A great deal of debate occurred, both at trial and on appeal, as to whether the deaths could be described as SIDS, which ultimately was a hollow and unnecessary discussion [1,2]. The legally relevant question was not whether the label SIDS could be attached as a reasonable possibility but whether the prosecution had proved homicide beyond reasonable doubt. The statistical evidence risked obscuring that distinction.

4. Could such an injustice occur now?

A disaster is sometimes best understood as the result of a cascade of smaller failures. In the Clark case, the failures were not minor but they were cumulative. Significant flaws are seen at each stage: the lack of integrity in Dr Williams' investigation, Professor Meadow's willingness to stray beyond his expertise, the clearly unscientific nature of his statistical evidence, the failure to exclude that evidence from consideration by the jury and the consequential incorporation of the prosecutor's fallacy into the trial, culminating in the refusal of the judges on the first appeal to recognise that the jury were likely influenced by clearly unscientific evidence.

If the question of whether such an injustice could occur now is asked in the specific context of the facts of the Sally Clark case, the answer is likely not. That, however, answers only the narrowest version of the question. The more important issue is whether an analogous failure could occur in another case, where serious weaknesses in the evidential foundation are overlooked until the case has already acquired momentum before a jury.

4.1. Scientific developments in SIDS/SUDI since Clark

Scientific understanding of SIDS/SUDI has considerably developed since 1999. Current research reflects a broader biological account of sudden unexplained infant death, including autonomic mechanisms, cardiorespiratory regulation, genetic susceptibility, metabolic vulnerability, infection, inflammation and bacterial-toxin hypotheses [23]. Some studies have used molecular and multiomic methods to investigate infection or neuroinflammation in deaths classified as SIDS, including processes which may not be apparent on standard autopsy [24]. Others have used genomic or cardiac-focused postmortem testing to investigate whether some infants may have had an underlying vulnerability, including variants associated with cardiac, neurological, systemic or immunological pathways [23,25,26]. These findings require caution, particularly where variants remain of uncertain significance. They nevertheless support the proposition that recurrent unexplained infant death cannot be assessed only by reference to the findings identified from conventional postmortem investigation. Metabolomic work adds a similar point, suggesting that some cases classified as SIDS may contain biological signals which require methods not routinely available or not routinely applied [27].

Because SIDS is a residual category, its evidential meaning depends on the scope and quality of the investigation used to exclude other causes. As the range of potentially relevant biological mechanisms has

expanded, and as investigative methods have become more sophisticated, the label "unexplained" has become increasingly dependent on the biological possibilities considered, the tests performed and the methods available at the time [23].

In *Clark*, the failure lay at the evidential stage rather than at the stage of detection. The postmortem examination of Clark's second child had identified the presence of *S. aureus*, but that finding was not initially disclosed and therefore did not form part of the evidential assessment placed before the jury [2]. Current literature gives that omission further scientific context, since *S. aureus* is commonly reported in the upper and lower respiratory tract of SIDS cases, and staphylococcal enterotoxin genes have been found more frequently in SIDS cases than in comparison infants [23,28]. In that setting, bacterial colonisation could not be dismissed as incidental without evaluation, because toxin-producing strains represent one possible biological mechanism in an otherwise unexplained infant death [28].

The significance of these developments is that unexplained infant death is now understood as a more heterogeneous and investigation-dependent category, less readily reducible to the binary framing of coincidence or homicide that influenced the reasoning in Clark. The contemporary literature gives no basis for assuming that every unexplained infant death conceals a natural cause. It does, however, make that binary difficult to justify where biological investigation remains incomplete, contested, inadequately disclosed or limited by the methods available.

4.2. Post-Clark reform: Law Commission, Reed and the Criminal Practice Directions

As a consequence of the Sally Clark and other controversial cases involving expert evidence, the Law Commission for England and Wales, the body responsible for proposing law reform in this jurisdiction, initially prepared a consultation paper in 2009 [29]. In that paper, the Law Commission made a number of suggestions, including the adoption of an enhanced reliability-based approach to expert evidence in criminal proceedings. The consultation paper drew comparison with the Daubert approach in United States litigation, under which courts act as gatekeepers of scientific expert evidence by permitting pre-trial challenge to evidence whose methodological reliability is in question [29,30]. Developed within the US Federal Rules of Evidence, and most prominently applied in civil litigation, Daubert provides a procedural mechanism through which parties may challenge the methodological reliability of expert scientific opinion before it reaches the jury [30–32]. Its application in criminal proceedings is less consistent and more constrained than in civil litigation but the existence of such a mechanism reflects a more structured approach to the admissibility of expert evidence. The comparison is important because it shows that the Law Commission was concerned not only with the formal qualification of experts but with whether the reasoning process underlying expert evidence could be scrutinised before it reached the jury.

After publication of the Law Commission consultation paper, the Court of Appeal again considered expert evidence, on this occasion in the context of DNA evidence, in *R v Reed and Garmson* [33]. At paragraph 111 of the judgment, the Court described the Law Commission's discussion of a Daubert-style enhanced test of admissibility as "valuable" but confirmed that such an approach was not appropriate in English courts. Rather, the Court confirmed the practice of the trial judge forming an assessment as to whether the evidence was sufficiently reliable to be admitted and considered by the jury. As the Law Commission later observed in its 2011 report, the Court of Appeal offered no explanation as to how a judge was to assess whether such evidence was sufficiently reliable [34].

In its 2011 report, the Law Commission proposed a statutory test for the admissibility of expert evidence, involving criteria similar to those used in reliability-based admissibility analysis under Daubert [34]. It also indicated subsidiary requirements in relation to qualification of experts and the process of admission of expert evidence prior to trial. These were set out in a draft Bill. In its discussion of the existing common law approach, the Law Commission described it as "laissez-faire". It quoted from a speech by Lord Justice Leveson to the Forensic Science Society, in which he stated:

> "It is, in my opinion, perfectly clear that expert evidence of doubtful reliability may be admitted too freely with insufficient explanation of the basis for reaching specific conclusions, be challenged too weakly by the opposing advocate and be accepted too readily by the Judge or Jury at the end of the Trial. In that regard, therefore, the law of England and Wales is not satisfactory and reform is undoubtedly required" [35].

In addition to making detailed recommendations for reform of the law, the Law Commission also stated that there should be "*a more critical approach on the part of some Judges to the evidence placed before them*". It also recommended appropriate training for all judges and lawyers involved in criminal proceedings on the assessment of evidential reliability, particularly where evidence is scientific in nature [34]. The observations of the Law Commission therefore highlighted that the problem exposed by the Clark case was not merely the conduct of one expert in one case. It was recognised by the Commission as part of a wider structural problem concerning the admission, challenge and evaluation of expert evidence in criminal proceedings.

The Government did not respond to the Law Commission report until November 2013 [35]. The Ministry of Justice substantially rejected the recommendations of the Law Commission, principally on grounds of cost. The Government response stated: *"The impact assessment published with the Law Commission's recommendations indicates that application of the new test would involve additional Pre-Trial Hearings, with concomitant additional costs, but without sufficient reliable predictable savings to compensate for these costs."*

The response referred to a lack of robust evidence as to the scale of the problem. It did not engage with the effect of a miscarriage of justice upon a given individual such as Sally Clark, nor upon the loss of credibility to the criminal justice system from the exposure of injustice. The Government did recommend that the issue be considered by the Criminal Procedure Rule Committee with the intention of considering amendment of the Criminal Procedure Rules. The only specific amendment to the Rules proposed in the Government response was that judges should be provided with "*more information about the expert evidence it is proposed to adduce*" [35].

Although this specific recommendation has not been enacted, the Criminal Practice Directions of 2015 dealt with expert evidence at section 19 [36]. Referring to the Law Commission Report and the Government's rejection of a statutory test, the Practice Direction stated in paragraph 19A.3: *"The common law, therefore, remains the source of the criteria by reference to which the Court must assess the admissibility and weight of such evidence; and [this Practice Direction] lists those matters which an expert report must deal with so that the Court can conduct an adequate such assessment".*

The Practice Direction further stated: *"Nothing at common law precludes assessment by the court of the reliability of an expert opinion by reference to substantially similar factors to those the Law Commission recommended as conditions of admissibility and Courts are encouraged actively to enquire into such factors".*

The Practice Direction then listed factors which could indicate potential problems with expert evidence, such as inadequate data, unjustified inference and experts straying beyond their specialist fields. The 2015 Practice Direction was superseded by the 2023 Practice Direction [37]. Section 7 of the 2023 Practice Direction addresses the admissibility of expert evidence by reference to relevance, necessity, competence and sufficient reliability. It then identifies 26 factors relevant to the assessment of reliability, including quality of data, methodology, unjustifiable assumptions, bias, and adverse previous findings or criticism [37].

4.3. Reliability criteria and the limits of judicial scrutiny

Superficially, the position is therefore more structured than it was at the time of Clark. However, the 2023 Practice Direction does not by itself eliminate the underlying vulnerability. It provides a list of factors by which reliability may be assessed but its effectiveness depends upon judges and advocates being able to identify the methodological flaw in the first place. A checklist cannot operate mechanically and requires those applying it to recognise the methodological significance of the evidence, particularly where an apparently authoritative opinion depends on assumptions, extrapolations or comparisons that have not been made explicit or properly justified.

Applying the current criteria to evidence of the kind presented in Clark illustrates the point. Its relevance would be doubtful if the figure were used, expressly or implicitly, as a proxy for the probability of innocence. Its necessity would also be questionable because, absent a valid comparative framework, it added little beyond the obvious proposition that two unexplained infant deaths in one family are rare. The competence requirement would raise a further issue if the witness's expertise lay in paediatrics rather than statistics or epidemiology. The reliability requirement would also be engaged by the unjustified assumption of independence, the limitations of the underlying data and the absence of comparative probabilistic analysis.

It could therefore be concluded with confidence that the evidence ought to be excluded in accordance with the current Practice Direction [37]. However, it does not automatically follow that it would be excluded. The Practice Direction requires the court to be astute to identify potential flaws in expert evidence. Both the Court of Appeal's retrospective view that exclusion would have been the "wisest" course and the Practice Direction's prospective requirement that courts be astute place significant weight on judicial discernment. The difficulty is that discernment in this context depends on a technical knowledge base that judges and advocates may not possess.

4.4. The Daubert comparison

Daubert remains useful as a comparison because it illustrates the importance of a structured pre-trial mechanism through which the methodological reliability of expert scientific evidence can be tested before it acquires forensic force before the jury [30,34]. Its usefulness is nevertheless limited by the fact that it is not invoked automatically and depends upon a party identifying the methodological problem and making the appropriate challenge. Its effectiveness therefore depends on the same anterior condition as the English approach: someone must first recognise that the expert evidence is not merely contestable but methodologically defective.

The English system in 1999 had no equivalent structured framework. The Criminal Practice Directions 2023 have since introduced reliability criteria that give judges a basis on which to scrutinise expert evidence more rigorously [37]. However, they do not provide a dedicated pre-trial hearing mechanism of the kind that Daubert contemplates. The Sally Clark case identifies the consequences of an unstructured approach. The defence challenged the statistical evidence at trial and raised the prosecutor's fallacy on first appeal but without the analytical foundation needed to expose the more fundamental methodological errors. Those errors were not publicly and authoritatively identified until the Royal Statistical Society intervened after the first appeal had already failed [8,9].

A structured pre-trial mechanism for methodological scrutiny would have provided a more appropriate forum for that analysis but the utility of any such mechanism depends upon the parties being equipped to use it effectively. The underlying vulnerability is therefore common to both systems; where the analytical sophistication required to recognise a statistical error is absent, procedural frameworks of whatever kind are unlikely to compensate for it.

4.5. Cannings and the requirement for extraneous cogent evidence

In 2004, the appeal of Angela Cannings was allowed by the Court of Appeal [38]. Cannings had lost three children in infancy and was convicted of murdering two of them at a trial in which Professor Meadow also gave evidence for the prosecution. The case involved disputed expert evidence concerning whether the deaths were natural or inflicted, in circumstances where there was no independent evidence that Cannings had harmed her children [38]. In allowing the appeal, the Court indicated that such cases should not be pursued where there was "*serious disagreement between distinguished and reputable experts*" unless the prosecution could point to "*extraneous cogent evidence*" outside the disputed expert opinion itself, such as clear evidence of abuse, an admission or other independent facts strongly supporting homicide [38].

In Sally Clark's case, there was also no such extraneous cogent evidence. There was no direct evidence that she had ever abused her children. All assessments of her indicated that she was a caring and careful mother [1,2]. The evidence called on behalf of the prosecution, even before the postmortem findings were discredited, could still be described as being largely theoretical in nature; that is, it drew significant conclusions by inference from the postmortem findings. In circumstances where reputable expert evidence was consistent with innocence, proceeding with a prosecution would appear difficult to understand in the context of proof beyond reasonable doubt. The obligation on prosecutors may, however, be blurred by the adversarial process. That difficulty was illustrated by the prosecution's attempt in Sally Clark's second appeal to uphold the conviction, notwithstanding the very obvious problems with its case at that stage. The prosecution only dropped its opposition to the appeal when the Court indicated that its position was hopeless.

4.6. Practical safeguards for analogous failures

The remaining question is what safeguards would reduce the risk of an analogous failure in a different case. The Clark case points to the need for earlier and more disciplined scrutiny of the evidential foundations of expert opinion, particularly where the opinion depends on probability, recurrence risk, incidence, prevalence or statistical inference. In such cases, professional eminence or experience cannot be treated as a substitute for an explicit and testable reasoning process.

That may require closer attention to the data source, the assumptions made, the inferential steps taken, the limits of the expert's discipline and the way in which any probability is being used in relation to competing explanations. In more complex cases, independent statistical or epidemiological input may be necessary before the evidence is placed before the jury. This is particularly important where the opinion depends on probabilistic reasoning, because the legal significance of such evidence may depend less on the figure itself than on whether its assumptions, reference class and inferential role have been properly understood.

Relevant competence and training are therefore critical to the effective application of any reliability-based admissibility framework, a point supported by Dahlman's empirical study of whether training in Bayesian thinking could help lawyers recognise probabilistic fallacies in legal evidence. Bayesian thinking, in this context, refers to reasoning about how evidence changes the probability of a hypothesis by comparing how likely the evidence is under competing explanations [39]. In Dahlman's study, participants who had received 5 h of classroom training in Bayesian reasoning were more accurate in recognising probabilistic fallacies than those who had received no such training. Overall accuracy increased from 63.7% in the control group to 82.9% in the trained group. However, the study also showed that some fallacies remained difficult to explain, particularly inversion fallacies such as the prosecutor's fallacy. Training is therefore necessary but insufficient on its own, where the evidential problem requires recognition of a technical probabilistic error [39].

The Clark case also illustrates the possible value, and the limits, of court-appointed expertise. An alternative approach might involve the court appointing its own expert, either to report on an independent basis or to act as an assessor. In Clark, the defence was able to call evidence from Professor David, a distinguished paediatrician whose involvement arose from his instruction by another court considering care arrangements for Sally Clark's third child. The fact that his evidence had not been elicited by either party was noted as significant in the judgment on the second appeal [2]. However, appointment by the court would not itself have solved the problem. Had such a system for appointing court-appointed experts existed in 1999, then Professor Meadow may well have appeared to be an appropriate appointment because of his professional standing in paediatrics. The critical question would remain whether the appointed expert had the correct expertise for the specific evidential issue before the court.

The same point applies to expert competence more generally. The difficulty exposed by *Robb* and *Clark* is that an expert may be qualified in a broad field while lacking the specific methodological expertise required for the evidential question before the court. Specialist training in the giving of expert evidence may assist, as may the disclosure of previous adverse judicial comment required by the Criminal Practice Directions 2023, but neither removes the need to consider whether the expert is qualified to give the particular opinion advanced [37].

4.7. Demeanour and the appearance of guilt

The jury system is itself subject to ongoing debate. As noted earlier, it is an ancient approach, rooted in the historical division between the judge's responsibility for law and the jury's responsibility for fact-finding. Proponents of the system consider that judgement by peers is a bastion of liberty going back to Magna Carta in 1215. Critics of it point to its cost and delay and the difficulty of juries reaching informed decisions in what might be described as complex cases [12–14]. One aspect of that difficulty concerns the jury's assessment of credibility. Within adversarial systems, the fact-finder has traditionally been regarded as having an advantage because they see and hear the witness. Wellborn notes that procedural rules concerning live testimony, confrontation and appellate deference have all been influenced by the assumed value of observing witness demeanour. He also refers to Mattox v United States, where the United States Supreme Court described confrontation as allowing the jury to look at the witness and judge, from demeanour upon the stand and the manner in which evidence is given, whether the witness is worthy of belief [40].

There is, however, substantial reason to treat that assumption with caution. The empirical literature on deception and nonverbal behaviour suggests that demeanour is a poor guide to truthfulness, that popular assumptions about body language may distort credibility assessment and that unaided human lie detection is only slightly better than chance [41–45]. Research on grief and resilience adds a further caution, since emotional responses to bereavement vary considerably and apparent composure should not be treated as inherently abnormal or suspicious [46].

In Sally Clark's case, some contemporary and later accounts of the trial commented that her demeanour remained calm almost throughout. She was a qualified solicitor and the daughter of a senior police officer, and her composure may therefore have reflected restraint, shock or professional self-control. The concern is that it may nevertheless have been capable of being read adversely, giving the impression of emotional coldness [47].

It is in this limited sense that the case invites comparison with the fate of the anti-hero Meursault in Albert Camus' classic novel The Outsider. Meursault was sentenced to death after a trial which involved, amongst other things, the prosecution producing evidence of his lack of emotion at his mother's funeral [48]. Camus summarised the book in one sentence: *"In our society any man who does not cry at his mother's funeral is liable to be condemned to death"*. It is useful because it captures a familiar forensic danger, namely that a person's apparent lack of expected emotion may be treated as if it reveals character, guilt or moral abnormality.

That danger is intensified where it forms part of a wider evidential failure. In Clark, the jury were not considering demeanour in a vacuum. They were considering it against a background of disputed medical evidence, confident expert opinion, statistical rarity and a prosecution case into which the prosecutor's fallacy had been incorporated. In that context, apparent composure could take on a significance far beyond its true evidential value and assist in giving apparent coherence to a case which already appeared to carry scientific force.

4.8. The remaining vulnerability

The answer to the question posed in the beginning of section 4 is therefore necessarily qualified. If the question is whether Clark would recur in precisely the same form, the answer is likely not. The errors made are now notorious. The one in 73 million statistic has been comprehensively criticised. If any expert sought to introduce statistical evidence which they were not qualified to give, the Court would, with overwhelming likelihood, refuse to admit it. The reasoning associated with Meadow's Law has also been strongly discredited and the Criminal Practice Directions now contain reliability criteria requiring disclosure of previous adverse judicial criticism of an expert [6,10,17,37]. Those developments make an exact repetition of Clark improbable. They do not, however, remove the structural risk exposed by the case. An analogous failure could still occur where expert evidence appears authoritative, where the reasoning depends on specialist assumptions not readily visible to lawyers or jurors, where probabilistic or scientific evidence is not properly tested before trial and where impressions of credibility or demeanour reinforce an apparently coherent prosecution narrative. The necessary safeguard is therefore a legal process capable of identifying when the evidential basis of expert opinion requires specialist scrutiny before it is placed before the jury, rather than reliance on any single rule, training requirement or checklist.

5. Conclusion

Ultimately, the most significant consideration relates to the relationship between expert opinion and the evidential basis on which it rests. The confidence with which some experts, certainly at the time, were prepared to express conclusions with potentially devastating effect was starkly illustrated by the conduct of another well-known paediatrician, Professor David Southall. Having watched a documentary about Sally Clark on television prior to her second appeal, which cast doubt on her guilt, he contacted the police to state that, in his view, the deaths had been caused not by Sally Clark but by her husband, Stephen Clark. He later gave evidence to that effect at a case conference in relation to the care and custody of the Clarks' third child. He was disciplined by the General Medical Council [49,50].

Professor Southall's conduct was not an isolated failing. It reflects a broader structural difficulty considered throughout this article in the relationship between expert integrity and the way expert evidence is engaged within the legal process. The Clark case demonstrates the risk inherent in a system in which confident expert opinion can be advanced, admitted and relied upon without sufficiently structured scrutiny of the assumptions on which it rests. The prevention of similar injustices depends on both expert integrity and legal process. Expert witnesses must remain within the limits of their expertise and courts, in turn, must have mechanisms capable of testing the evidential foundation of expert opinion before it acquires forensic force before a jury.

CRediT authorship contribution statement

Charles Feeny: Conceptualization, Writing – original draft. **Carolina Stamboulid:** Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R v Clark, EWCA Crim 54, 2000.
[2] R v Clark, EWCA Crim 1020, 2003.
[3] G.Y.C. Cheung, J.S. Bae, M. Otto, Pathogenicity and virulence of Staphylococcus aureus, Virulence 12 (1) (2021) 547–569.
[4] C. Dyer, Falsely convicted Sally Clark dies suddenly, BMJ 334 (7594) (2007) 602–603.
[5] S. Gaines, D. Pallister, Sally Clark's death accidental, coroner rules, Guardian (2007 7 November 2007).
[6] R. Hill, Multiple sudden infant deaths - coincidence or beyond coincidence? Paediatr. Perinat. Epidemiol. 18 (5) (2004) 320–326.
[7] R. Nobles, D. Schiff, Misleading statistics within criminal trials, Significance 2 (1) (2005) 17–19.
[8] Royal Statistical Society, Royal Statistical Society Concerned by Issues Raised in the Sally Clark Case, 2001. London.
[9] Royal Statistical Society, Letter from the President to the Lord Chancellor Regarding the Use of Statistical Evidence in the Sally Clark Case, 2002. London.
[10] M.D. Freeman, A.M. Rossignol, M.L. Hand, Forensic Epidemiology: a systematic approach to probabilistic determinations in disputed matters, J. Forensic Legal Med. 15 (5) (2008) 281–290.
[11] P. Dawid, Statistics on trial, Significance 2 (1) (2005) 6–8.
[12] J. Hodgson, The English Legal Heritage, Oyez Publishing, 1979.
[13] T. Plucknett, A Concise History of the Common Law, Butterworth & Co, 1956.
[14] G. Rivlin, Understanding the Law, Oxford, 2004.
[15] R v Robb, 93 Cr App 161, 1991.
[16] R. Meadow, ABC of child abuse. Epidemiology, BMJ 298 (6675) (1989) 727–730.
[17] R. Hill, Reflections on the cot death cases, Significance 2 (1) (2005) 13–16.
[18] P.B.P.S.P. Fleming, C. Bacon, P.J. Berry, Sudden unexpected death in infancy. The CESDI SUDI Studies 1993-1996, The Stationery Office, London, 2000.
[19] P. Dawid, Bayes' theorem and weighing evidence by juries, in: R. Swinburne (Ed.), Bayes' Theorem. 113, Proceedings of the British Academy, 2002, pp. 71–90.
[20] R v Doheny, Adams, 1 Cr App R 369, 1997.
[21] R.W. Byard, Sudden infant death syndrome: definitions, in: J.R. Duncan, R. W. Byard (Eds.), SIDS Sudden Infant and Early Childhood Death: the Past, the Present and the Future, University of Adelaide Press, Adelaide (AU), 2018.
[22] H.F. Krous, J.B. Beckwith, R.W. Byard, T.O. Rognum, T. Bajanowski, T. Corey, et al., Sudden infant death syndrome and unclassified sudden infant deaths: a definitional and diagnostic approach, Pediatrics 114 (1) (2004) 234–238.
[23] P.N. Goldwater, Current SIDS research: time to resolve conflicting research hypotheses and collaborate, Pediatr. Res. 94 (4) (2023) 1273–1277.
[24] P.S. Ramachandran, B.W. Okaty, M. Riehs, A. Wapniarski, D. Hershey, H. Harb, et al., Multiomic analysis of neuroinflammation and occult infection in sudden infant death syndrome, JAMA Neurol. 81 (3) (2024) 240–247.
[25] A.M. Bard, L.V. Clark, E. Cosgun, K.A. Aldinger, A. Timms, L.A. Quina, et al., Known pathogenic gene variants and new candidates detected in sudden unexpected infant death using whole genome sequencing, Am. J. Med. Genet. 194 (11) (2024) e63596.
[26] F. Cazzato, M. Coll, S. Grassi, A. Fernàndez-Falgueras, L. Nogué-Navarro, A. Iglesias, et al., Investigating cardiac genetic background in sudden infant death syndrome (SIDS), Int. J. Leg. Med. 138 (6) (2024) 2229–2237.
[27] C.M. Aldridge, K.L. Keene, C.A. Normeshie, J.C. Mychaleckyj, F.R. Hauck, Metabolomic profiles of infants classified as sudden infant death syndrome: a case-control analysis, EBioMedicine 111 (2025) 105484.
[28] A.R. Highet, P.N. Goldwater, Staphylococcal enterotoxin genes are common in Staphylococcus aureus intestinal flora in Sudden Infant Death Syndrome(SIDS) and live comparison infants, FEMS Immunol. Med. Microbiol. 57 (2) (2009) 151–155.
[29] Law Commission, The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales, Law Commission, London, 2009.
[30] Daubert v Merrell Dow Pharmaceuticals Inc., 509 US 579, 1993.
[31] Federal Rules of Evidence, Rule 702 (2023).
[32] Federal Judicial Center, Reference manual on scientific evidence. Federal Judicial Center, third ed., 2011 18 June 2026. Washington, DC.
[33] R. v Reed, Reed, Garmson, EWCA Crim 2698, 2009.
[34] Law Commission, Expert Evidence in Criminal Proceedings in England and Wales (Law Com No 325), 2011.
[35] The Government's Response to the Law Commission Report, Expert Evidence in Criminal Proceedings in England and Wales, Law Com No 325, Ministry of Justice, London, 2013.
[36] Criminal Practice Directions (2015). London: Judiciary of England and Wales; 2015.
[37] Criminal Practice Directions (2023). London: Judiciary of England and Wales; 2023.
[38] R v Cannings, EWCA Crim 1, 2004.
[39] C. Dahlman, Training legal fact-finders to recognize probabilistic fallacies, Law Probab. Risk 24 (1) (2025).
[40] O.G. Wellborn III., Demeanor, Cornell Law Rev. 76 (5) (1991) 1075–1105.
[41] C.F. Bond, B.M. DePaulo, Accuracy of deception judgments, Pers. Soc. Psychol. Rev. 10 (3) (2006) 214–234.
[42] V. Denault, Misconceptions about nonverbal cues to deception: a covert threat to the justice System? Front. Psychol. 11 (2020), 2020.
[43] V. Denault, N.E. Dunbar, Credibility assessment and deception detection in courtrooms: hazards and challenges for scholars and legal practitioners, in: T. Docan-Morgan (Ed.), The Palgrave Handbook of Deceptive Communication, Springer International Publishing, Cham, 2019, pp. 915–935.
[44] V. Denault, C. Leclerc, V. Talwar, The use of nonverbal communication when assessing witness credibility: a view from the bench, Psychiatr. Psychol. Law 31 (1) (2024) 97–120.
[45] M. Hartwig, Jr CF. Bond, Why do lie-catchers fail? A lens model meta-analysis of human lie judgments, Psychol. Bull. 137 (4) (2011) 643–659.
[46] G.A. Bonanno, Loss, trauma, and human resilience: have we underestimated the human capacity to thrive after extremely aversive events? Am. Psychol. 59 (1) (2004) 20–28.
[47] J. Batt, Stolen Innocence: a Mother's Fight for Justice: the Story of Sally Clark, Random House, London, 2005, p. 491.
[48] A. Camus, The Outsider, Penguin Books, London, 1989.
[49] C. Dyer, GMC hears case against paediatrician in Sally Clark trial, Br. Med. J. 330 (7506) (2005) 1463.
[50] C. Dyer, Professor Roy Meadow struck off, Br. Med. J. 331 (7510) (2005) 177.